

SEQUENCE LISTING

(1) GENERAL INFORMATION

- 5 (i) APPLICANT: Lincoln, Steve
Klinger, Tod M.
Au-Young, Janice
10 Tang, Y. Tom
Goold, Richard
Akerblom, Ingrid E.
Seilhamer, Jeffrey J.
Hawkins, Phillip R.
Murry, Lynn E.
15 Delegeane, Angelo M.
Levine, Wendy B.
Hillman, Jennifer L.
Goli, Surya K.
Altus, Christina M.
20 Bandman, Olga
LaBrie, Samuel T.
Shah, Purvi
- 25 (ii) TITLE OF THE INVENTION: Database for Storage and Analysis of Full
Length Sequences
- (iii) NUMBER OF SEQUENCES: 10
- 30 (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Bozicevic & Reed, L.L.P.
(B) STREET: 285 Hamilton Ave.
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: U.S.
35 (F) ZIP: 94301
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
40 (B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5
- (vi) CURRENT APPLICATION DATA:
45 (A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Filed Herewith
- (vii) PRIOR APPLICATION DATA:
50 (A) APPLICATION NUMBER: 08/282,955
(B) FILING DATE: July 29, 1995
(A) APPLICATION NUMBER: 08/187,530
(B) FILING DATE: January 27, 1994
55 (A) APPLICATION NUMBER: 08/179,873
(B) FILING DATE: January 11, 1994
(A) APPLICATION NUMBER: 08/100,523

(B) FILING DATE: August 3, 1993

(A) APPLICATION NUMBER: 07/977,780
(B) FILING DATE: November 19, 1992

(A) APPLICATION NUMBER: 07/916,491
(B) FILING DATE: July 17, 1992

(A) APPLICATION NUMBER: 08/289,822
(B) FILING DATE: August 12, 1994

(A) APPLICATION NUMBER: 08/581,240
(B) FILING DATE: December 29, 1995

(A) APPLICATION NUMBER: 08/657,697
(B) FILING DATE: May 29, 1996

(A) APPLICATION NUMBER: 08/747,547
(B) FILING DATE: November 12, 1996

(A) APPLICATION NUMBER: 08/712,710
(B) FILING DATE: September 12, 1996

(A) APPLICATION NUMBER: 08/744,026
(B) FILING DATE: November 5, 1996

(A) APPLICATION NUMBER: 08/786,999
(B) FILING DATE: January 23, 1997

(A) APPLICATION NUMBER: 08/822,262
(B) FILING DATE: March 20, 1997

(A) APPLICATION NUMBER: 08/000,000 (as yet unassigned)
(B) FILING DATE: October 16, 1997

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Francis, Carol L.
(B) REGISTRATION NUMBER: 36,513
(C) REFERENCE/DOCKET NUMBER: 6514/069001

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 650-327-3400
(B) TELEFAX: 650-327-3231

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 373 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: Pancreatic Cancer

(B) CLONE: PANC1A

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5 GTACGGAGGT GAGGTTTGTN ACCGCGATTC TAAGAGGTGG GCTTTTAGTC CCTCCAGACC 60
 TCGGCTTTAG TGCTGTCTCC GCTTTTYYTTT CACCTTCACA GAGGTTCTGT TCTTCCTAAA 120
 AGAAGGTTTT ATTGGGAGGT AAAGGTCAAT GCGTAGGGGT AGAGTAAGAT GTCTTATGGT 180
 GAAATTRAAG GTAAATTCTT GGGACCTAGA GAAGAAGTAA CGAGTGAGCC ACGCTGTAAA 240
 AAATTGAAGT CAACCACAGA GTCGTATGTT TTTACAATC ATAGTAATGC TGATTTTCAC 300
 10 AGNATCCAAG AGAAAACCTG AAATGATTGG GTCCCTGTGN NCATCATTGA TGTCAGAGGA 360
 CATAGTTATT TGC 373

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 321 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: Pancreatic Cancer

(B) CLONE: PANC1B

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

30 GTGAGGTTTG TTACCNCGAT TCTGAGAGGT GGGCTTTTAG TCCCTCCAGA CCTCGGCTTT 60
 AGTGCTGTCT CCGMTTCTT TACACCTTCA CAGAGATGTC TTATGGTGAA ATTGAAGGTA 120
 AATTCTTGGG ACCTAGWGAA GAAGTAACGA GTGAGCCACG CTGTAAAAAA TTGAAGTCAA 180
 CCACAGAGTC GTATGTTTTT CACAATCATA GTAATGCTGA TTTTCACAGW ATCCAAGAGA 240
 AAACCTGGAAA TGATTTGGGT CCCTGTGACC ATCATTNATG TCAGAGGNCA TAGTTAATTT 300
 35 GCAGGAGANC AAAAAATCAAA A 321

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 528 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: NEUTGMT01

(B) CLONE: 637393

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGACAGACT GTGAATTTGG ATATATTTAC AGGCTGGCTC AGGACTATCT GCAGTGCGTC 60
 CTACAGATAC CACAACCTGG ATCAGGTCCA AGCAAAACGT CCAGAGTGCT ACAAATGTT 120
 55 GCGTTCTCAG TCCAAAAAGA AGTGAAAAAG AATCTGAAGT CATGCTTGGA CAATGTTAAT 180
 GTTGTGTCCG TAGACACTGC CAGAACTACTA TTCAACCAAG TGATGGAAAA GGAGTTTGAA 240
 GACGACATCA TTAACCTGGG AAGAATTGTA ACCATATTTG CATTTGAAGG TATTCTCATC 300

AAGAACTTC TACGACAGCA AATTGCCCCG GATGTGGATA CCTATAAGGA GATTTTCATAT 360
 TTTGTTGCGG AGTTCATAAT GAATAACACA GGAGAATGGA TAAGGCAAAA CGGAGGCTGG 420
 GAAAATGGCT TTGTAAAGAA GTTTGAACCT AAATCTGGCT GGATGACTTT TCTAGAAGTT 480
 ACAGGAAAGA TCTGTGAAAT GCTATCTCTC CTGAAGCAAT ACTGTTGA 528

5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 405 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

- 15 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

20 GTCCAAATCA CTCATTGTTT GTGAAAGCTG AGCTCACAGC AAAACAAGCC ACCATGAAGC 60
 TGTCGGTGTG TCTCCTGCTG GTCACGCTGG CCCTCTGCTG CTACCAGGCC AATGCCGAGT 120
 TCTGCCCAGC TCTTGTCTCT GAGCTGTTAG ACTTCTTCTT CATTAGTGAA CCTCTGTTCA 180
 AGTTAAGTCT TGCCAAATTT GATGCCCCCTC CGGAAGCTGT TGCAGCCAAG TTAGGAGTGA 240
 AGAGATGCAC GGATCAGATG TCCCTTCAGA AACGAAGCCT CATTGCGGAA GTCCTGGTGA 300
 25 AAATATTGAA GAAATGTAGT GTGTGACATG TAAAAACTTT CATCCTGGTT TCCACTGTCT 360
 TTCAATGACA CCCTGATCTT CACTGCAGAA TGTAAGGTT TCAAC 405

30 (2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 495 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GATCCTTGCC ACCGCGGACT GAACACCGAC AGCAGCAGCC TCACCATGAA GTTGCTGATG 60
 GTCCTCATGC TGGCGGCCCT CTCCCAGCAC TGCTACGCAG GCTCTGGCTG CCCCTTATTG 120
 GAGAATGTGA TTTCCAAGAC AATCAATCCA CAAGTGCTCTA AGACTGAATA CAAAGAACTT 180
 45 CTTCAAGAGT TCATAGACGA CAATGCCACT ACAAATGCCA TAGATGAATT GAAGGAATGT 240
 TTTCTTAACC AAACGGATGA AACTCTGAGC AATGTTGAGG TGTTTATGCA ATTAATATAT 300
 GACAGCAGTC TTTGTGATTT ATTTTAACTT TCTGCAAGAC CTTTGGCTCA CAGAACTGCA 360
 GGGTATGGTG AGAAACCAAC TACGGATTGC TGCAAAACCAC ACCTTCTCTT TCTTATGTCT 420
 TTTTACTACA AACTACAAGA CAATTGTTGA AACCTGCTAT ACATGTTTAT TTTAATAAAT 480
 50 TGATGGCAAA AACTG 495

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 1143 base pairs
 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE: Consensus

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

	GGCCCCGCCG	CGCCCCGGCGC	GCCCCGCCGCC	CGGGGGGATG	TCTTACAAAC	CGAACTTGGC	60
	CGCGCACATG	CCCGCCGCCG	CCCTCAACGC	CGCTGGGAGT	GTCCACTCGC	CTTCCACCAG	120
	CATGGCAACG	TCTTCACAGT	ACCGCCAGCT	GCTCAGTGAC	TACGGGGCCAC	CGTCCCTAGG	180
15	CTACACCCAG	GGAACCTGGGA	ACAGCCAGGT	GCCCCAAAGC	AAATACGCGG	AGCTGCTGGC	240
	CATCATTGAA	GAGCTGGGGA	AGGAGATCAG	ACCCACGTAC	GCAGGGAGCA	AGAGTGCCAT	300
	GGAGAGGCTG	AAGCGCGGCA	TCATTACACG	TAGAGGACTG	GTTTCGGGAGT	TCTTGGCAGA	360
	AACGGAACGG	AATGCCAGAT	CCTAGCTGCC	TTGTTGGTTT	TGAAGGATTT	CCATCTTTTT	420
	ACAAGATGAG	AAGTTACAGT	TCATCTCCCC	TGTTTCAGATG	AAACCCCTTG	TTTCAAATG	480
20	GTTACAGTTT	CGTTTTTCCT	CCCATGGTTC	ACTTGGCTCT	GAACCTACAG	TCTCAAAGAT	540
	TGAGAAAAGA	TTTTGCAGTT	AATTAGGATT	TGCATTTTAA	GTAGTTAGGA	ACTGCCCAGG	600
	TTTTTTTTGT	TTTTTAAGCA	TTGATTTAAA	AGATGCACGG	AAAGTTATCT	TACAGCAAAC	660
	TGTAGTTTGC	CTCCAAGACA	CCATTGTCTC	CCTTTAATCT	TCTCTTTTGT	ATACATTTGT	720
	TACCCATGGT	GTTCTTTGTT	CCTTTTCATA	AGCTAATACC	ACTGTAGGGA	TTTTGTTTTG	780
25	AACGCATATT	GACAGCACGC	TTTACTTAGT	AGCCGGTTCC	CATTTGCCAT	ACAATGTAGG	840
	TTCTGCTTAA	TGTAACCTCT	TTTTTGCTTA	AGCATTTGCA	TGACTATTAG	TGCTTCAAAG	900
	TCAATTTTAA	AAAATGCACA	AGTTATAAAT	ACAGAAGAAA	GAGCAACCCA	CCAAACCTAA	960
	CAAGGACCCC	CGAACACTTT	CATACTAAGA	CTGTAAGTAG	ATCTCAGTTC	TGCGTTTATT	1020
	GTAAGTTGAT	AAAAACATCT	GGAAGAAAAT	GAATAAACT	GTTTGCATCT	TTGTATGTAT	1080
30	TTATTACTTG	ATGTAATAAA	GCTTATTTTC	ATTAACAATT	TGTATTAAAA	TGTGGGTTCC	1140
	TTG						1143

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 871 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

50	ACCTGCTGGC	CCCTGGACAC	CTCTGTCACC	ATGTGGTTCC	TGGTTCTGTG	CCTCGCCCTG	60
	TCCCTGGGGG	GGACTGGTGC	TGCGCCCCCG	ATTCAGTCCC	GGATTGTGGG	AGGCTGGGAG	120
	TGTGAGCAGC	ATTCCCAGCC	CTGGCAGGCG	GCACTGGTCA	TGGAAAACGA	ATTGTTCTGC	180
	TCGGGCGTCC	TGGTGCATCC	GCAGTGGGTG	CTGTACGCCG	CACACTGTTT	CCAGAACTCC	240
	TACACCATCG	GGCTGGGCCT	GCACAGTCTT	GAGGCCGACC	AAGAGCCAGG	GAGCCAGATG	300
55	GTGGAGGCCA	GCCTCTCCGT	ACGGCACCCA	GAGTACAACA	GACCCTTGCT	CGCTAACGAC	360
	CTCATGNTCA	TCAAGTTGGA	CGAATCCGTG	TCCGAGTCTG	ACAACATCCG	GAGNATCAGC	420
	ATTGNTTCGC	AGTGCCCTAC	CGCGGGGAAC	TTTTGCCTCG	TTTCTGGCTG	GGGTCTGCTG	480

GCGAACGGCA GAATGCCTAC CGTGCTGCAG TCGGTGAACG TGTCGGTGGT GTCTGAGGAG 540
 GTCTGCAGTA AGCTCTATGA CCCGCTGTAC CACCCAGCA TGTTCCTGCGC CGGCGGAGGG 600
 CAAGACCAGA AGGACTCCTG CAACGGTGAC TCTGGGGGGC CCCTGATCTG CAACGGGTAC 660
 TTGCAGGGCC TTGTGTCTTT CGGAAAAGCC CCGTGTGGCC AAGTTGGCGT GCCAGGTGTC 720
 5 TACACCAACC TCTGCAAATT CACTGAGTGG ATAGAGAAAA CCGTCCAGGC CAGTTAACTC 780
 TGGGGACTGG GAACCCATGA AATTGACCCC CAAATACATC CTGCGGAAGG AATTCAGGAA 840
 TATCTGTTCC CAGCCCCCTCC TCCCTCAGGC C 871

10 (2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1428 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Colnnot01
 20 (B) CLONE: 609476

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

25 GGCCGGGTTG GGGGTGTGCG ATTGTGTGGG ACGGTCTGGG GCAGCCCAGC AGCGGCTGAC 60
 CCTCTGCCTG CGGGGAAGGG AGTCGCCAGG CGGCCGTCAT GGCGGTGTCG GAGAGCCAGC 120
 TCAAGAAAAT GGTGTCCAAG TACAAATACA GAGACCTAAC TGTACGTGAA ACTGTCAATG 180
 TTATTACTCT ATACAAAGAT CTCAAACCTG TTTTGGATTG ATATGTTTTT AACGATGGCA 240
 GTTCCAGGGA ACTAATGAAC CTCACTGGAA CAATCCCTGT GCCTTATAGA GGTAATACAT 300
 ACAATATTCC AATATGCCTA TGGCTACTGG ACACATACCC ATATAATCCC CCTATCTGTT 360
 30 TTGTTAAGCC TACTAGTTCA ATGACTATTA AAACAGGAAA GCATGTTGAT GCAAATGGGA 420
 AGATATATCT TCCTTATCTA CATGAATGGA AACACCCACA GTCAGACTTG TTGGGGCTTA 480
 TTCAGGTCAT GATTGTGGTA TTTGGAGATG AACCTCCAGT CTTCTCTCGT CCTATTTCTG 540
 CATCCTATCC GCCATACCAG GCAACGGGGC CACCAAATAC TTCCTACATG CCAGGCATGC 600
 CAGGTGGAAT CTCTCCATAC CCATCCGGAT ACCCTCCCAA TCCCAGTGGT TACCCAGGCT 660
 35 GTCCTTACCC ACCTGGTGGT CCATATCCTG CCACAACAAG TTCTCAGTAC CTTCTCAGC 720
 CTCCTGTGAC CACTGTTGGT CCCAGTAGGG ATGGCACAAT CAGCGAGGAC ACCATCCGAG 780
 CCTCTCTCAT CTCTGCGGTC AGTGACAAAC TGAGATGGCG GATGAAGGAG GAAATGGATC 840
 GTGCCCAGGC AGAGCTCAAT GCCTTGAAAC GAACAGAAGA AGACCTGAAA AAGGGTCACC 900
 AGAAACTGGA AGAGATGGTT ACCCGTTTAG ATCAAGAAGT AGCCGAGGTT GATAAAAACA 960
 40 TAGAACTTTT GAAAAAGAAG GATGAAGAAC TCAGTTCTGC TCTGGAAAAA ATGGAAAATC 1020
 AGTCTGAAAA CAATGATATC GATGAAGTTA TCATTCCCAC AGCTCCCTTA TACAAACAGA 1080
 TCCTGAATCT GTATGCAGAA GAAAACGCTA TTGAAGACAC TATCTTTTAC TTGGGAGAAG 1140
 CCTTGAGAAG GGGCGTGATA GACCTGGATG TCTTCTGAA GCATGTACGT CTTCTGTCCC 1200
 GTAAACAGTT CCAGCTGAGG GCACTAATGC AAAAAGCAAG AAAGACTGCC GGTCTCAGTG 1260
 45 ACCTCTACTG ACTTCTCTGA TACCAGCTGG AGGTTGAGCT CTTCTTAAAG TATTCTTCTC 1320
 TTCCTTTTAT CAGTAGGTGC CCAGAATAAG TTATTGCAGT TTATCATTCA MGTGTAAAAT 1380
 ATTTTGAATC AATAATATAT TTTCTGTTTT CTTTGGTAA AGATGGAT 1428

50 (2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 667 base pairs
 (B) TYPE: nucleic acid
 55 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: THPINOT03

(B) CLONE: 2446131

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

5
 GGCTGAGCGG CCCCAGAGCC AACCCCGAG GAGCGGCCGG CTGGCGTCCG CCGCGCCCAG 60
 GAGTTGGGGA TGTCCTACAA ACCCATCGCC CCTGCTCCCA GCAKCAACCC TGGCTCCAGC 120
 ACCCCTGGGC CGGGCACCCC GGTCCCTACA GGAAGCGTCC CGTCGCCGTC GGGCTCAGTG 180
 CCAGGAGCCG GCGCTCCTTT CAGACCGCTG TTTAACGACT TTGGACCGCC TTCCATGGGC 240
 10 TACGTGCAGG CGATGAAGCC ACCCGGCGCC CAGGGCTCCC AGAGCACCTA CACGGACCTG 300
 CTGTCTAGTCA TAGAGGAGAT GGGCAAAGAG ATCCGGCCTA CCTATGCTGG CAGCAAGAGC 360
 GCCATGGAGC GCCTGAAGAG AGGTATCATC CATGCCCGGG CCCTAGTCAG AGAGTGCCTG 420
 GCAGAGACAG AGCGGAACGC CCGCACGTAA CAGGAAGCGC CTCGGCCTCA GCGTCTGGAC 480
 CTATCCGGCC ACTGCAGAGC ACCCGCTTCT CCCTGGCCTT CATCCCGAGT TGCCTAACC 540
 15 ATCCTGGGCT TCCTGTCCTG TGTCCCTTGG TGGGTCCCTT CCAGGAACCA AGGAGTGGCC 600
 CTCCAGGTGG CAGCACTAAG GACACCCCCC CACAACAAGA GTTAGCAGCG AGGTCCCCAT 660
 GAGTCCC 667

(2) INFORMATION FOR SEQ ID NO:10: -

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 397 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

30 CTGAACTCTA CCTGGTGACC AGGGACCAGG ACCTTTATAA GGTGGAAGGC TTGATGTCCT 60
 CCCCAGACTC AGCTCCTGGT GAAGCTCCCA GCCATCAGCC ATGAGGGTCT TGTATCTCCT 120
 CTTCTCGTTC CTCTTCATAT TCCTGATGCC TCTTCCAGGT GTTTTTGGTG GTATAGGCGA 180
 TCCCGTTACC TGCCTTAAGA GTGGAGCCAT ATGTCATCCA GTCTTTTGCC CTAGAAGGTA 240
 TAAACNAATT GGCACCTGTG GTCTCCCTGG AACAAAATGC TGNAAAAAGC CATGAGGAGG 300
 35 CCAAGAAGCT GCTGTGGCNG ANGCGATTG AGAAAGGGCT CCCTCATCAG AGANGTGCGA 360
 CATGTAAACC AAATTAACT ATGGTGTCCA ANGATAN 397